

Scalable Generation of Scale-free Graphs

Peter Sanders, Christian Schulz

Karlsruhe Institute of Technology (KIT), 76128 Karlsruhe, Germany
`{sanders,christian.schulz}@kit.edu`

Abstract. We explain how massive instances of scale-free graphs following the Barabasi-Albert model can be generated very quickly in an embarrassingly parallel way. This makes this popular model available for studying big data graph problems. As a demonstration, we generated a Petaedge graph in less than an hour.

1 Introduction

Scale-free graphs with a power-law degree distribution seem to be ubiquitous in complex network analysis. In order to study such networks and the algorithms to analyze them, we need simple models for generating complex networks with user-definable parameters. Barabasi and Albert [3] define the model that is perhaps most widely used because of its simplicity and intuitive definition: We start with an arbitrary seed network consisting of nodes $0..n_0 - 1$ ($a..b$ is used as a shorthand for $\{a, \dots, b\}$ here). Nodes $i \in n_0..n - 1$ are added one at a time. They randomly connect to d neighbors using *preferential attachment*, i.e., the probability to connect to node $j \leq i$ is chosen proportionally to the degree of j . The seed graph, n_0 , d , and n are parameters defining the graph family.

With the recent interest in big data, the inherently sequential definition of these graphs has become a problem however, and other, less natural models have been considered for generating very large networks, e.g., for the well known Graph500 benchmark [1].

2 Our Algorithm

Our starting point is the fast, simple, and elegant sequential algorithm by Batagelj and Brandes [4]. For simplicity of exposition, we first consider the most simple situation with an empty seed graph

and where self-loops and parallel edges are allowed. See Section 4 for generalizations.

We use an empty seed graph ($n_0 = 0$). A generalization only requires a number of straight forward index transformations. Batagelj and Brandes' algorithm generates one edge at a time and writes it into an edge array $E[0..2dn - 1]$ where positions $2i$ and $2i + 1$ store the node IDs of the end points of edge i . We have $E[2i] = \lfloor i/d \rfloor$. The central observation is that one gets the right probability distribution for the other end point by uniformly sampling edges rather than sampling dynamically weighted nodes, i.e., $E[2i + 1]$ is simply set to $E[x]$ where x is chosen uniformly and (pseudo)randomly from $0..2i$.

The idea behind the parallel algorithm is very simple – compute edge i independently of all other edges and without even accessing the array E . On the first glance, this sounds paradoxical because there *are* dependencies and accessing E is the whole point behind Batagelj and Brandes' algorithm. This paradox is resolved by the idea to recompute any entry of E that is needed for edge i . However, this solution raises two concerns. First, doesn't recomputation increase the amount of work in an unacceptable way? Second, how do you reproduce random behavior?

The first concern is resolved by observing that we have a 50 % chance of looking at an even position which is easy to reproduce. In the other 50 % of the cases, we have to look at further positions of E . Overall, the expected number of positions of E considered is bounded by $\sum_{i \geq 0} 2^{-i} = 2$ – compared to Batagelj and Brandes' algorithm, we compute around twice as many random numbers but in exchange save most of the expensive memory accesses.

What saves the situation with respect to reproducing random behavior is that practical computer programs usually do not use true randomness but only pseudorandomness, i.e., if you know the state of a pseudorandom number generator, you can reproduce its results deterministically. This is a nontrivial issue in a parallel setting but becomes easy using another trick. We use a hash function that maps the array position to a pseudorandom number. For a graph generator this approach has the additional benefit that the graph only depends on the hash function but not on the number of processors used to compute the graph. The pseudo-code below summarizes the resulting

```

function generateEdge( $i$ )                                     // generate  $i$ -th edge
 $r := 2i + 1$ 
repeat  $r := h(r)$  until  $r$  is even
return ( $\lfloor i/d \rfloor, \lfloor r/2d \rfloor$ )

```

Fig. 1. Pseudocode of generateEdge.

algorithm where $h(r)$ is a hash function mapping an integer r to a random number in $0..r - 1$.

Note that this simple setting allows us a number of interesting approaches. Using dynamic load balancing, handing out batches of edge IDs, we can use cheap heterogeneous cloud resources. We do not even have to store the graph if the analysis uses a distributed streaming algorithm that processes the edges immediately.

3 Experiments

As a simple example, we have implemented an algorithm finding the degrees of the first $100K$ nodes of a graph with $n = 10^{13}$ nodes and $m = 100n$ edges on 16 384 cores of the SuperMUC supercomputer¹ (see Figure 2). As a hash function we use the CRC32 instruction available since SSE 4.2 twice with random seeds to obtain 64 bits of pseudo-random data (see Figure 3). Preliminary experiments indicated that this is slightly faster than using a simple hash function. In comparison, Batagelj and Brandes algorithm is about 60 % faster than our algorithm on a single core but slower than any parallel run. We are about 16 times faster than the parallel RMAT generator [1] for a graph with $50 \cdot 10^9$ edges. Alam et al. [2] report 123 s for generating a graph with $50 \cdot 10^9$ edges on 768 processors. They use an algorithm explicitly tracing dependencies which leads to massive amounts of fine grained communication. We are about 36 times faster on a machine with slower cores and using 64 bit node IDs rather than 32 bits. Meyer and Penschuk [6] can generate a graph with the same parameters on a single node with 16 cores and a GPU

¹ The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (LRZ, www.lrz.de).

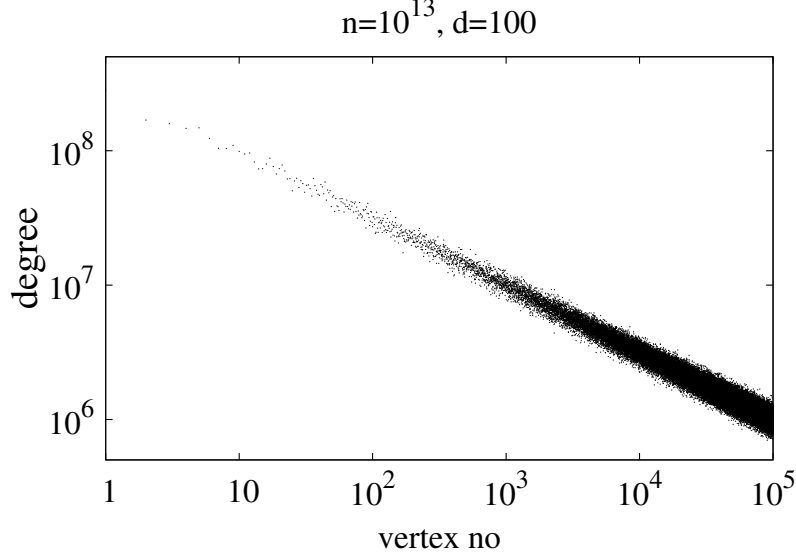


Fig. 2. The degrees of the first 100K nodes of a graph with $n = 10^{13}$ nodes and $m = 100n$ edges computed on 16 384 cores of the SuperMUC supercomputer.

using an ingenious external memory algorithm and 6 SSDs in 489 s. Our streaming generator using just the 16 cores needs about 154 s. For smaller graphs, using 32 bits instead of 64 bits everywhere in our code saves about 65% of running time. The largest graph we have generated is 20 000 times larger than the largest Barabasi-Albert graph we have seen reported.

4 Generalizations

A seed graph with n_0 nodes and m_0 edges is incorporated by stopping the repeat loop in the edge generation function when $r < m_0$. For $r < m_0$ the node ID is precomputed. Otherwise, we compute $\lfloor (r - m_0)/d \rfloor + n_0$. Self-loops can be avoided by initializing r to the first edge of the current node. To avoid parallel edges, we store the target node IDs generated for a node in a hashtable of size $\leq d$ and reject duplicate target nodes. Now consider a situation with individual degrees d_i for each node. If about $\sum_i \log d_i$ bits fit into each

```

function hash_crc(x) {
    hash = _mm_crc32_u64(0, x);
    hash = hash << 32;
    hash += _mm_crc32_u64(1, x);
    return hash % x;
}

hashMultiplier=3141592653589793238LL;
function h_simple(x) { return (x*hashMultiplier)*(1.0/0xffffffffffffLL) * x; }

```

Fig. 3. Hash functions used in our implementations.

processor, we can replicate a succinct representation of a sparse bit vector [5] with a one bit for each first edge of a node. Then the target node of edge r can be computed in constant time using the operation $\text{rank}(r)$. For larger networks, we can defer the computation of node IDs and first only compute edge IDs². At the end, we can sort these edge IDs and merge the result with the prefix sum of the node degree sequence to obtain the node IDs. Any parallel sorting and merging algorithm can be used for this purpose.

² We would like to thank Ulrik Brandes for pointing this out.

References

1. Graph 500 Benchmark www.graph500.org/.
2. Maksudul Alam, Maleq Khan, and Madhav V Marathe. Distributed-memory parallel algorithms for generating massive scale-free networks using preferential attachment model. In *Proc. of the Int. Conference on High Performance Computing, Networking, Storage and Analysis*, page 91. ACM, 2013.
3. Albert-Laszlo Barabasi and Reka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
4. Vladimir Batagelj and Ulrik Brandes. Efficient generation of large random networks. *Physical Review E*, 71(3):036113, 2005.
5. Simon Gog and Matthias Petri. Optimized succinct data structures for massive data. *Software: Practice and Experience*, 44(11):1287–1314, 2014.
6. Ulrich Meyer and Manuel Penschuk. Generating massive scale free networks under resource constraints. In *Meeting on Algorithm Engineering & Experiments (ALENEX)*, pages 39–52. SIAM, 2016.